MAS223 Statistical Modelling and Inference

Exercises and Solutions

The exercises are grouped into sections, corresponding to chapters of the lecture notes. Within each section exercises are divided into warm-up questions, ordinary questions, and challenge questions. Note that there are no exercises accompanying Chapter 8.

The vast majority of exercises are ordinary questions. Ordinary questions will be used in homeworks and tutorials; they cover the material content of the course. Warm-up questions are typically easier, often nothing more than revision of relevant material from first year courses. Challenge questions are typically harder and test ingenuity.

This version of the exercises also contains solutions, which are written in blue. Solutions to challenge questions are not always included, hints may be given instead. Some of the solutions mention common pitfalls, written in red, which are mistakes that are (sometimes) easily made.

The solutions sometimes omit intermediate steps of basic calculations, which are left to the reader. For example, they may simply state $\int_0^x \lambda e^{-\lambda u} = 1 - e^{-\lambda u}$, and leave you to fill in the intermediate steps.

Contents

1 Univariate Distribution Theory

Warm-up Questions

1.1 Let X be a random variable taking values in $\{1, 2, 3\}$, with $\mathbb{P}[X = 1] = \mathbb{P}[X = 2] = 0.4$. Find $\mathbb{P}[X = 3]$, and calculate both $\mathbb{E}[X]$ and $\text{Var}[X]$.

Solution. Since $\mathbb{P}[X = 1] + \mathbb{P}[X = 2] + \mathbb{P}[X = 3] = 1$, we have $\mathbb{P}[X = 3] = 0.2$. With this, we can calculate

$$\mathbb{E}[X] = 1\mathbb{P}[X = 1] + 2\mathbb{P}[X = 2] + 3\mathbb{P}[X + 3] = 1.8$$
$$\mathbb{E}[X^2] = 1^2\mathbb{P}[X = 1] + 2^2\mathbb{P}[X = 2] + 3^2\mathbb{P}[X + 3] = 3.8$$

Using that $\text{Var}(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2$, we have $\text{Var}(X) = 0.56$.

1.2 Let Y be a random variable with probability density function (p.d.f.) $f(y)$ given by

$$f(y) = \begin{cases} y/2 & \text{for } 0 \le y < 2; \\ 0 & \text{otherwise.} \end{cases}$$

Find the probability that Y is between $\frac{1}{2}$ and 1. Calculate $\mathbb{E}[Y]$ and $\text{Var}[Y]$.

Solution. We have $\mathbb{P}[Y \in [\frac{1}{2}, 1]] = \int_{1/2}^{1} (y/2)\, dy = 3/16$. Similarly,

$$\mathbb{E}[Y] = \int_{-\infty}^{\infty} y f(y) = \int_0^2 y(y/2)\, dy = 4/3$$
$$\mathbb{E}[Y^2] = \int_0^2 (y^3/2)\, dy = 2$$

so $\text{Var}(Y) = \mathbb{E}[Y^2] - \mathbb{E}[Y]^2 = 2/9$.

Ordinary Questions

1.3 Define $F : \mathbb{R} \to [0, 1]$ by

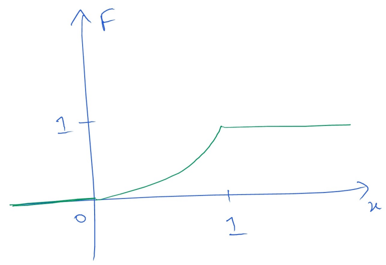
$$F(y) = \begin{cases} 0 & \text{for } y \le 0; \\ y^2 & \text{for } y \in (0, 1); \\ 1 & \text{for } y \ge 1. \end{cases}$$

(a) Sketch the function F, and check that it is a distribution function.

(b) If Y is a random variable with distribution function F, calculate the p.d.f. of Y.

Solution.

(a) Sketch of F should look like



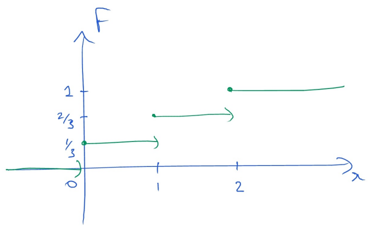
From the graph, F is continuous and (non-strictly) increasing. We have $F(x) = 0$ for all $x \le 0$, so $\lim_{x\to-\infty} F(x) = 0$. Similarly, $F(x) = 1$ for all $x \ge 1$, so $\lim_{x\to\infty} F(x) = 1$. Hence, F satisfies all the properties of a distribution function.

(b) We have $f(y) = F'(y)$, so treating each case in turn,

$$f(y) = \begin{cases} 0 & \text{for } y \le 0; \\ 2y & \text{for } y \in (0,1); \\ 0 & \text{for } y \ge 1. \end{cases}$$

1.4 Let X be a discrete random variable, taking values in $\{0,1,2\}$, where $\mathbb{P}[X = n] = \frac{1}{3}$ for $n \in \{0,1,2\}$. Sketch the distribution function $F_X : \mathbb{R} \to \mathbb{R}$.

Solution. Sketch should look like



Pitfall: The graph of F is not continuous; it jumps at $0,1,2$, and is otherwise constant.

1.5 Define $f : \mathbb{R} \to [0,1]$ by

$$f(x) = \begin{cases} 0 & \text{for } x < 0; \\ e^{-x} & \text{for } x \ge 0. \end{cases}$$

(a) Show that f is a probability density function.

(b) Find the corresponding distribution function and evaluate $\mathbb{P}[1 < X < 2]$.

Solution.

(a) Clearly $f(x) \ge 0$ for all x, and

$$\int_{-\infty}^{\infty} f(x)\,dx = \int_0^\infty e^{-x}\,dx = \left[-e^{-x}\right]_0^\infty = 1,$$

so f is a probability density function.

(b) We need to calculate $F(x) = \mathbb{P}[X \leq x] = \int_{-\infty}^{x} f(u)\,du$. For $x \leq 0$ we have $F(x) = \int_{-\infty}^{x} 0\,dx = 0$. For $x \geq 0$, we have $F(x) = \int_{-\infty}^{0} 0\,du + \int_{0}^{x} e^{-u}, du = 0 + (1 - e^{-x})$. Thus,

$$F(x) = \begin{cases} 0 & \text{for } x \leq 0; \\ 1 - e^{-x} & \text{for } x \geq 0. \end{cases}$$

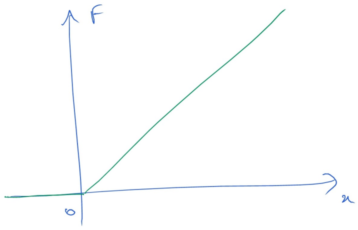
Hence, $\mathbb{P}[1 < X < 2] = \mathbb{P}[X < 2] - \mathbb{P}[X \leq 1] = \mathbb{P}[X \leq 2] - \mathbb{P}[X \leq 1] = F(2) - F(1) = e^{-1} - e^{-2}$.

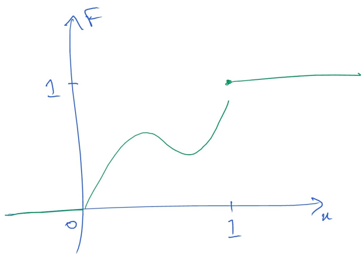
1.6 Sketch graphs of each of the following two functions, and explain why each of them is not a distribution function.

(a) $F(x) = \begin{cases} 0 & \text{for } x \leq 0; \\ x & \text{for } x > 0. \end{cases}$

(b) $F(x) = \begin{cases} 0 & \text{for } x < 0; \\ x + \frac{1}{4}\sin 2\pi x & \text{for } 0 \leq x < 1; \\ 1 & \text{for } x \geq 1. \end{cases}$

Solution. Sketches should look like





For (a), $F(x) > 1$ for $x > 1$, so F does not stay between 0 and 1. For (b), for $x \in [0, 1)$ we have $F'(x) = f(x) = 1 + \frac{2\pi}{4}\cos(2\pi x)$, which is negative at, for example, $x = \frac{1}{2}$, so (as is clear from the graph) F is not an increasing function.

1.7 Let $k \in \mathbb{R}$ and define $f : \mathbb{R} \to \mathbb{R}$ by

$$f(x) = \begin{cases} k(x - x^2) & \text{for } x \in (0, 1), \\ 0 & \text{otherwise.} \end{cases}$$

Find the value of k for which $f(x)$ is a probability density function, and calculate the probability that X is greater than $\frac{1}{2}$.

Solution. We need $f(x) \geq 0$ for all x, so we need $k \geq 0$. Also, we need

$$1 = \int_{-\infty}^{\infty} f(x)\,dx = k\int_{0}^{1} x - x^2\,dx = k\left[\frac{x^2}{2} - \frac{x^3}{3}\right]_0^1 = \frac{k}{6}.$$

So, $k = 6$. Therefore,

$$\mathbb{P}[X \geq \tfrac{1}{2}] = \int_{\frac{1}{2}}^{\infty} f(x)\,dx = \int_{\frac{1}{2}}^{1} 6(x - x^2)\,dx = 6\left[\frac{x^2}{2} - \frac{x^3}{3}\right]_{\frac{1}{2}}^{1} = \frac{1}{2}.$$

1.8 The probability density function $f(x)$ is given by

$$f(x) = \begin{cases} 1+x & \text{for } -1 \leq x < 0; \\ 1-x & \text{for } 0 \leq x < 1; \\ 0 & \text{otherwise.} \end{cases}$$

Find the corresponding distribution function $F(x)$ for all real x.

Solution. We have $F(-1) = \mathbb{P}[X \leq -1] = P[X < -1] = 0$. So $F(x) = 0$ for all $x \leq -1$. If $x \in [-1, 0]$ then

$$F(x) = \int_{-\infty}^{x} f(u)\, du = F(-1) + \int_{-1}^{x} (1+u)\, du = 0 + \frac{(x+1)^2}{2}.$$

Now, for $x \in [0, 1]$, we have

$$F(x) = \int_{-\infty}^{x} f(u)\, du = F(0) + \int_{0}^{x} (1-u)\, du = \frac{1}{2} + x - \frac{x^2}{2} = \frac{1+2x-x^2}{2}.$$

Pitfall: Forgetting the F(0) results in missing out a factor of $\frac{1}{2}$. The factor $F(0)$ covers the part of the integral from $\int_{-\infty}^{0}$, and from the case $x \in [-1, 0]$ we know that $F(0) = \frac{1}{2}$. Note that in the case of $x \in [-1, 0]$ the equivalent term was $F(-1)$ and was equal to 0.

Therefore, we have $F(1) = 1$. Since F is increasing and must stay between 0 and 1, we have $F(x) = 1$ for all $x \geq 1$.

Thus the distribution function $F(x)$ is

$$F(x) = \begin{cases} 0, & \text{for } x < -1 \\ \frac{(x+1)^2}{2}, & \text{for } -1 \leq x < 0 \\ \frac{1+2x-x^2}{2}, & \text{for } 0 \leq x < 1 \\ 1, & \text{for } x \geq 1 \end{cases}$$

1.9 Let

$$F(x) = \frac{e^x}{1+e^x} \quad \text{for all real } x.$$

(a) Show that F is a distribution function, and find the corresponding p.d.f. f.

(b) Show that $f(-x) = f(x)$.

(c) If X is a random variable with this distribution, evaluate $\mathbb{P}[|X| > 2]$.

Solution.

(a) Since $e^x \to 0$ as $x \to \infty$, we have

$$\lim_{x\to-\infty} \frac{e^x}{1+e^x} = \frac{0}{1+0} = 0$$

$$\lim_{x\to\infty} \frac{e^x}{1+e^x} = \lim_{x\to\infty} \frac{1}{e^{-x}+1} = \frac{1}{0+1} = 1$$

Pitfall: It does not make sense to use that $\lim_{x\to\infty} e^x = \infty$ and then incorrectly claim that $\frac{\infty}{1+\infty} = 1$.

Using the quotient rule, the derivative of f, the corresponding p.d.f. , is

$$f(x) = \frac{e^x(1+e^x) - e^x e^x}{(1+e^x)^2} = \frac{e^x}{(1+e^x)^2}.$$

Therefore, $f(x) > 0$ for all $x \in \mathbb{R}$, so F is an increasing function.

Since $x \mapsto e^x$ is continuous, F is a composition of sums and (non-zero) divisions of continuous functions; therefore F is continuous.

Hence, F satisfies all the properties of a distribution function.

(b) $f(-x) = \frac{e^{-x}}{(1+e^{-x})^2} = \frac{e^{2x}e^{-x}}{e^{2x}(1+e^{-x})^2} = \frac{e^x}{(e^x+1)^2} = f(x)$.

(c) We have

$$\begin{aligned}\mathbb{P}[|X| > 2] &= \mathbb{P}[X < -2] + \mathbb{P}[X > 2] \\ &= \mathbb{P}[X \le -2] + (1 - \mathbb{P}[X \le 2]) \\ &= 1 + F(-2) - F(2) \\ &= 1 + \tfrac{e^{-2}}{1+e^{-2}} - \tfrac{e^2}{1+e^2} \approx 0.238.\end{aligned}$$

Note that here we used $\mathbb{P}[X < -2] = \mathbb{P}[X \le -2]$, which holds because F is continuous.

1.10 The discrete random variable X has the probability function

$$\mathbb{P}[X = x] = \begin{cases} \frac{1}{x(x+1)} & \text{for } x \in \mathbb{N} \\ 0 & \text{otherwise.} \end{cases}$$

(a) Use the partial fractions of $\frac{1}{x(x+1)}$ to show that $\mathbb{P}[X \le x] = 1 - \frac{1}{x+1}$, for all $x \in \mathbb{N}$.

(b) Write down the distribution function $F(x)$ of X, for $x \in \mathbb{R}$. Sketch its graph. What are the values of $F(2)$ and $F(\frac{3}{2})$?

(c) Evaluate $\mathbb{P}[10 \le X \le 20]$.

(d) Is $\mathbb{E}[X]$ defined? If so, what is $\mathbb{E}[X]$? If not, why not?

Solution.

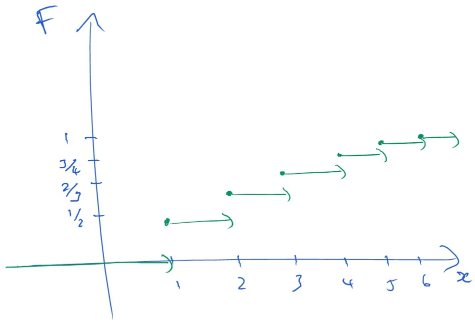
(a) Using partial fractions we obtain the identity $\frac{1}{x(x+1)} = \frac{1}{x} - \frac{1}{x+1}$, provided $x \ne 0, -1$. Hence, if $x \in \mathbb{N}$,

$$\begin{aligned}\mathbb{P}[X \le x] &= \sum_{i=1}^{x} \mathbb{P}[X = x] \\ &= \left(\frac{1}{1} - \frac{1}{2}\right) + \left(\frac{1}{2} - \frac{1}{3}\right) + \left(\frac{1}{3} - \frac{1}{4}\right) + \ldots + \left(\frac{1}{x} - \frac{1}{x+1}\right) \\ &= 1 - \frac{1}{x+1}.\end{aligned}$$

(b) The distribution function is

$$F(u) = \begin{cases} 1 - \frac{1}{x+1}, & \text{for } u \in [x, x+1) \text{ where } x \in \mathbb{N} \\ 0, & \text{for } u < 1 \end{cases}$$

which looks like



$F(2) = 1 - \frac{1}{3} = \frac{2}{3}$ and $F(\frac{3}{2}) = F(1) = \frac{1}{2}$.

Pitfall: The graph of F is not continuous. The formula obtained in part (a) is only valid for $x \in \mathbb{N}$, and not for all $x \in \mathbb{R}$. Since X is a discrete random variable, its distribution function jumps at the points where X takes values (i.e. at $x \in \mathbb{N}$) and is constant in between those points.

(c) Since X is discrete,

$$\mathbb{P}[10 \le X \le 20] = \mathbb{P}[X \le 20] - \mathbb{P}[X \le 9] = F(20) - F(9) = \tfrac{11}{210}.$$

Pitfall: X is discrete, and $\mathbb{P}[X = 10] > 0$. So it's $F(20) - F(9)$, and not $F(20) - F(10)$.

(d) Since X is discrete, $\mathbb{E}[X]$ is defined if and only if $\sum_x x\mathbb{P}[X = x]$ converges. In our case, this sum is equal to

$$\sum_{x=1}^{\infty} x \frac{1}{x(x+1)} = \sum_{x=1}^{\infty} \frac{1}{x+1}$$

which diverges.

To see that the sum diverges, we can write it as

$$\left(\frac{1}{2}\right) + \left(\frac{1}{3} + \frac{1}{4}\right) + \left(\frac{1}{5} + \frac{1}{6} + \frac{1}{7} + \frac{1}{8}\right) + \dots$$

and note that each bracketed term is at least $\frac{1}{2}$; of course $\sum_{x=1}^{\infty} \frac{1}{2} = \infty$.

Challenge Questions

1.11 For which values of $r \in [0, \infty)$ is $\int_1^\infty x^{-r}\,dx$ finite? Give an example of a continuous random variable for which $\mathbb{E}[X]$ is defined but $\mathbb{E}[X^2]$ is not.

Hint. Use a p.d.f. of the form

$$f(x) = \begin{cases} Cx^{-r} & \text{for } x \ge 1 \\ 0 & \text{otherwise.} \end{cases}$$

where $C = \left(\int_1^\infty x^{-r}\,dx\right)^{-1}$. You'll need to find an appropriate value of r.

1.12 Let $A = \mathbb{Q} \cap [0,1]$. Since A is countable, let us write $A = (a_1, a_2, a_3, \ldots)$, where $a_n \neq a_m$ for $n \neq m$. Let $F : \mathbb{R} \to \mathbb{R}$ be given by

$$F(x) = \sum_{n=1}^{\infty} 2^{-n} \mathbb{1}_{\{a_n \le x\}}.$$

Here, $\mathbb{1}_{\{a_n \le x\}}$ is equal to 1 if $a_n \le x$, and equal to 0 otherwise.

(a) Show that F is a distribution function.

(b) Let X be a random variable with distribution function F. Show that $\mathbb{P}[X = x] = 0$ for all $x \in \mathbb{R} \setminus A$, and that $\mathbb{P}[X = a_n] = 2^{-n}$ for all $x \in A$.

Hint. Try thinking about what the graph of F would look like for a simpler set A, such as $A = (\frac{1}{4}, \frac{1}{2}, \frac{3}{4})$. Imagine gradually adding more elements of $\mathbb{Q}$ into A.

The point is that the graph of F jumps by 2^{-n} at each $x = a_n \in A$. It doesn't jump at each $x \notin A$, but each such x has infinitely many (mostly small) jumps within the neighbourhood $(x - \epsilon, x + \epsilon)$ for any $\epsilon > 0$.

Solution.

(a) Since $\sum_{n=1}^{\infty} 2^{-n} = 1$, and $F(x)$ picks up only a subset of the terms 2^{-n}, we have $F(x) \in [0,1]$ for all x. Since $F(x) = 0$ for $x < 0$ and $F(x) = \sum_{n=1}^{\infty} 2^{-n} = 1$ for $x > 1$, we have $\lim_{x \to -\infty} F(x) = 0$ and $\lim_{x \to \infty} F(x) = 1$.

For $x < y$ we have $\mathbb{1}_{\{a_n \le x\}} \le \mathbb{1}_{\{a_n \le y\}}$ so as $F(x) \le F(y)$.

It remains to show that F is right continuous with left limits. We have

$$\begin{aligned} 0 \le F(x + \epsilon) - F(x) &= \sum_{n=1}^{\infty} 2^{-n} \mathbb{1}\{x < a_n \le x + \epsilon\} \\ &\le \sum_{n=N(\epsilon)}^{\infty} 2^{-n} \end{aligned} \tag{1.1}$$

where $N(\epsilon) = \inf\{n \in \mathbb{N} \;:\; x < a_n \le x + \epsilon\}$. As $\epsilon \to 0$, $N(\epsilon) \to \infty$, hence (1.1) $\to 0$, which means that

$$\lim_{\epsilon \to 0} F(x + \epsilon) = F(x).$$

That is, F is right continuous. To show that F has left limits, we note that

$$\begin{aligned} 0 \le \sum_{n=1}^{\infty} 2^{-n} \mathbb{1}_{\{a_n < x\}} - F(x - \epsilon) &= \sum_{n=1}^{\infty} 2^{-n} \mathbb{1}_{\{x - \epsilon < a_n < x\}} \\ &\le \sum_{n=N'(\epsilon)}^{\infty} 2^{-n} \end{aligned} \tag{1.2}$$

where $N'(\epsilon) = \inf\{n \in \mathbb{N} \;:\; x - \epsilon < a_n < x\}$. As $\epsilon \to 0$, $N(\epsilon) \to \infty$, hence (1.2) $\to 0$, which means that

$$\lim_{\epsilon \to 0} F(x - \epsilon) = \sum_{n=1}^{\infty} 2^{-n} \mathbb{1}_{\{a_n < x\}}.$$

That is, the left limit of F at x is given by

$$F(x-) = \sum_{n=1}^{\infty} 2^{-n} \mathbb{1}_{\{a_n < x\}}.$$

(b) From (a), for any x we have

$$\begin{aligned}\mathbb{P}[X = x] &= F(x) - F(x-) \\ &= \sum_{n=1}^{\infty} 2^{-n} \mathbb{1}_{\{a_n \le x\}} - \sum_{n=1}^{\infty} 2^{-n} \mathbb{1}_{\{a_n < x\}} \\ &= \sum_{n=1}^{\infty} 2^{-n} \left(\mathbb{1}_{\{a_n \le x\}} - \mathbb{1}_{\{a_n < x\}} \right).\end{aligned}$$

For $x \notin A$, for all n we have $\mathbb{1}\{a_n \le x\} = \mathbb{1}\{a_n < x\}$. Thus, $\mathbb{P}[X = x] = 0$. On the other hand, for $x = a_m \in A$ we have $\mathbb{1}\{a_n \le a_m\} = \mathbb{1}\{a_n < a_m\}$ for all $n \neq m$, but for the case $m = n$ we have $\mathbb{1}\{a_m \le a_n\} = 1$ and $\mathbb{1}\{a_m < a_n\} = 0$. Thus, $\mathbb{P}[X = a_m] = 2^{-m}$.

2 Standard Univariate Distributions

Warm-up Questions

2.1 (a) A standard fair dice is rolled 5 times. Let X be the number of sixes rolled. Which standard distribution (and which parameters) would you use to model X?

(b) A fair coin is flipped until the first head is shown. Let X be the total number of flips, including the final flip on which the first head appears. Which standard distribution (and which parameter) would you use to model X?

Solution.

(a) The binomial distribution, with parameters $n = 5$ and $p = \frac{1}{6}$.

(b) The geometric distribution, with parameter $p = \frac{1}{2}$.

Ordinary Questions

2.2 Let $\lambda > 0$. Write down the p.d.f. f of the random variable X, where $X \sim Exp(\lambda)$, and calculate its distribution function F. Hence, show that $\frac{f(t)}{1-F(t)}$ is constant for $t > 0$.

Solution. The p.d.f. of X is

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & \text{for } x > 0; \\ 0 & \text{otherwise.} \end{cases}$$

It's distribution function $F(x) = \int_{-\infty}^{x} f(u)\,du$ is clearly zero for $x \le 0$, and for $x > 0$ we have $\int_{-\infty}^{x} f(u)\,du = \int_{0}^{x} \lambda e^{-\lambda u}\,du = 1 - e^{-\lambda x}$. Therefore,

$$F(x) = \begin{cases} 1 - e^{-\lambda x} & \text{for } x > 0; \\ 0 & \text{otherwise.} \end{cases}$$

Hence, for $t > 0$ we have $\frac{f(t)}{1-F(t)} = \frac{\lambda e^{-\lambda t}}{e^{-\lambda t}} = \lambda$.

Pitfall: Don't forget the 'otherwise' case where $f(x) = 0$ or $F(x) = 0$. The same comment applies to many other questions.

2.3 Let $\lambda > 0$ and let X be a random variable with $Exp(\lambda)$ distribution. Let $Z = \lfloor X \rfloor$, that is let Z be X rounded down to the nearest integer. Show that Z is geometrically distributed with parameter $p = 1 - e^{-\lambda}$.

Solution. Since $X > 0$, we have $Z \in \{0, 1, 2, \ldots\}$, hence $\mathbb{P}[Z = z] = 0$ for all other z. For $n \in \{0, 1, 2, \ldots\}$ we have

$$\begin{aligned} \mathbb{P}[Z = n] &= \mathbb{P}[n \le X < n + 1] \\ &= \int_{n}^{n+1} \lambda e^{-\lambda x}\,dx \\ &= e^{-\lambda n} - e^{-\lambda(n+1)} \\ &= e^{-\lambda n}(1 - e^{-\lambda}) \\ &= (1 - p)^n p. \end{aligned}$$

which is the probability function of the geometric distribution.

2.4 Let $\mu \in \mathbb{R}$. Let X_1 and X_2 be independent random variables with distributions $N(\mu, 1)$ and $N(\mu, 4)$, respectively. Let T_1, T_2 and T_3 be defined by

$$T_1 = \frac{X_1 + X_2}{2}, \qquad T_2 = 2X_1 - X_2, \qquad T_3 = \frac{4X_1 + X_2}{5}.$$

Find the mean and variance of T_1, T_2 and T_3. Which of $\mathbb{E}[T_1]$, $\mathbb{E}[T_2]$ and $\mathbb{E}[T_3]$ would you prefer to use as an estimator of μ?

Solution. We have $\mathbb{E}[T_1] = \frac{1}{2}(\mathbb{E}[X_1] + \mathbb{E}[X_2]) = \mathbb{E}[T_2] = \mu$. Similarly, $\mathbb{E}[T_2] = \mathbb{E}[T_3] = \mu$, so all are unbiased when used as estimators of μ. We have

$$\mathrm{Var}(T_1) = (\tfrac{1}{2})^2 \left(\mathrm{Var}(X_1) + \mathrm{Cov}(X_1, X_2) + \mathrm{Var}(X_2)\right) = \tfrac{1}{4}(1 + 0 + 4) = \tfrac{5}{4},$$

and similarly $\mathrm{Var}(T_2) = 8$, $\mathrm{Var}(T_3) = \frac{4}{5}$.

On this information, we prefer $\mathbb{E}[T_3]$ as an estimator of μ, because T_3 has the smallest variance and so is likely to be closest to its mean.

2.5 Let X be a random variable with $Ga(\alpha, \beta)$ distribution.

(a) Let $k \in \mathbb{N}$. Show that

$$\mathbb{E}[X^k] = \frac{\alpha(\alpha+1)\cdots(\alpha+k-1)}{\beta^k}.$$

Hence, calculate $\mu = \mathbb{E}[X]$ and $\sigma^2 = \mathrm{Var}(X)$ and verify that these formulas match the ones given in lectures.

(b) Show that $\mathbb{E}\left[\left(\frac{X-\mu}{\sigma}\right)^3\right] = \frac{2}{\sqrt{\alpha}}$.

Solution.

(a) From the p.d.f. of the Gamma distribution, we have

$$\begin{aligned}
\mathbb{E}[X^k] &= \int_0^\infty x^k \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x}\, dx \\
&= \frac{\beta^\alpha}{\Gamma(\alpha)} \int_0^\infty x^{k+\alpha-1} e^{-\beta x}\, dx \\
&= \frac{\beta^\alpha}{\Gamma(\alpha)} \frac{\Gamma(k+\alpha)}{\beta^{k+\alpha}} \\
&= \frac{\Gamma(k+\alpha)}{\beta^k \Gamma(\alpha)} \\
&= \frac{(k+\alpha-1)(k+\alpha-2)\cdots(k+\alpha-k+1)(k+\alpha-k)\Gamma(\alpha)}{\beta^k \Gamma(\alpha)} \\
&= \frac{\alpha(\alpha+1)\cdots(\alpha+k-1)}{\beta^k}.
\end{aligned}$$

To deduce the second line from the to third line, we use Lemma 2.2 from lecture notes, and to deduce the fifth line from the fourth line we use Lemma 2.1, also from lecture notes.

Pitfall: It is true that $\Gamma(n) = (n-1)!$ for $n \in \mathbb{N}$, but if $n \notin \mathbb{N}$ then $(n-1)!$ does not make sense. For general $\alpha \in (1,\infty)$ we have $\Gamma(\alpha) = (\alpha-1)\Gamma(\alpha-1)$, which is what must be be used to deduce the fifth line above.

If $k = 1$ then $\mathbb{E}[X] = \frac{a}{b}$. If $k = 2$ then $\mathbb{E}[X^2] = \frac{\alpha(\alpha+1)}{\beta^2}$ and so the variance of X is $\mathrm{Var}(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2 = \frac{\alpha(\alpha+1)}{\beta^2} - \frac{\alpha^2}{\beta^2} = \frac{\alpha}{\beta^2}$.

(b) With $\mu = \mathbb{E}[X]$ and $\sigma^2 = \mathrm{Var}(X)$, we have

$$\begin{aligned}
\frac{\mathbb{E}[X-\mu]^3}{\sigma^3} &= \frac{\mathbb{E}[X^3 - 3X^2\mu + 3X\mu^2 - \mu^3]}{\sigma^3} \\
&= \frac{\mathbb{E}[X^3] - 3\mathbb{E}[X]\mathbb{E}[X^2] + 2\mathbb{E}[X]^3}{\sigma^3} \\
&= \frac{\frac{\alpha(\alpha+1)(\alpha+2)}{\beta^3} - \frac{3\alpha^2(\alpha+1)}{\beta^3} + \frac{2\alpha^3}{\beta^3}}{\sigma^3} \\
&= \frac{\alpha(\alpha^2 + 2\alpha + \alpha + 2 - 3\alpha^2 - 3\alpha + 2\alpha^2)/\beta^3}{\sqrt{\alpha^3}/\beta^3} \\
&= \frac{2}{\sqrt{\alpha}}.
\end{aligned}$$

2.6 (a) Using R, you can obtain a plot of, for example, the p.d.f. of a $Ga(3,2)$ random variable between 0 and 10 with the command

```
curve(dgamma(x,shape=3,scale=2),from=0,to=10)
```

Use R to investigate how the shape of the p.d.f. of a Gamma distribution varies with the different parameter values. In particular, fix a value of β, see how the shape changes as you vary α.

(b) Investigate the effect that changing parameters values has on the shape of the p.d.f. of the Beta distribution. To produce, for example, a plot of the p.d.f. of $Be(4,5)$, use

```
curve(dbeta(x,shape1=4,shape2=5),from=-1,to=2)
```

Solution.

(a) You should discover that decreasing α makes the p.d.f. appear more skewed (to the right). This makes it more likely that a sample of the random variable has a large value.

(b) You should discover that the parameter `shape1` (which we normally denote by α) controls the behaviour near $x = 0$, and `shape2` (that is, β), controls the behaviour near $x = 1$. In both case, the parameters can be tuned to cause (slow or fast) explosion to ∞, convergence to 1, and (slow or fast) convergence towards 0.

2.7 Suggest which standard discrete distributions (or combination of them) we should use to model the following situations.

(a) Organisms, independently, possess a given characteristic with probability p. A sample of k organisms with the characteristic is required. How many organisms will need to be tested to achieve this sample?

(b) In Texas Hold'em Poker, players make the best hand they can by combining two cards in their hand with five 'community' cards that are placed face up on the table. At the start of the game, a player can only see their own hand. The community cards are then turned over, one by one.

A player has two hearts in her hand. Three of the community cards have been turned over, and only one of them is a heart. How many hearts will appear in the remaining two community cards?

Find the probability of seeing $k = 0, 1, 2$ hearts.

Solution.

(a) We'll need to sample, with success probability p, until we achieve k successes. So we will need to test $N \sim NegBin(k, p)$ organisms to find our sample.

(b) There are a total of 52 cards, 13 of each of the four suits. Our player can see 5 cards, 3 of which are hearts. Therefore, the unknown cards consist of 47 cards, 10 of which are hearts. The number of hearts that will be drawn in the next two community cards is, therefore, a hypergeometric distribution with parameters $N = 47$ (population size), $k = 10$ (successes), $n = 2$ (trials). As a result (use e.g. `R`), $\mathbb{P}[X = 0] \approx 0.65$, $\mathbb{P}[X = 1] \approx 0.32$ and $\mathbb{P}[X = 2] = 0.03$.

2.8 Let X be a $N(0, 1)$ random variable. Use integration by parts to show that $\mathbb{E}[X^{n+2}] = \frac{1}{n+1}\mathbb{E}[X^n]$ for any $n = 0, 1, 2, \ldots$. Hence, show that

$$\mathbb{E}[X^n] = \begin{cases} 0 & \text{if } n \text{ is odd;} \\ \frac{1}{(1)(3)(5)\ldots(n-1)} & \text{if } n \text{ is even.} \end{cases}$$

Solution. Integrating by parts, for any $n \in \mathbb{N}$, gives

$$\begin{aligned} \mathbb{E}[X^n] &= \int_{-\infty}^{\infty} x^n \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \\ &= \left[\frac{x^{n+1}}{n+1} e^{-x^2/2}\right]_{-\infty}^{\infty} - \int_{-\infty}^{\infty} \frac{x^{n+1}}{n+1} \frac{1}{\sqrt{2\pi}} (-x) e^{-x^2/2} dx \\ &= 0 + \int_{-\infty}^{\infty} \frac{x^{n+2}}{n+1} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx = \frac{1}{n+1} \mathbb{E}[X^{n+2}] \end{aligned}$$

Since $\mathbb{E}[X] = 0$, induction gives that $\mathbb{E}[X^n] = 0$ for all odd n. Since $\mathbb{E}[X^2] = \mathrm{Var}(X) = 1$, induction gives that $\mathbb{E}[X^n] = \frac{1}{(1)(3)(5)\ldots(n-1)}$ for all even n.

2.9 Let $X \sim N(\mu, \sigma^2)$. Show that $\mathbb{E}[e^X] = e^{\mu + \frac{\sigma^2}{2}}$.

Solution. We need to calculate

$$\mathbb{E}[e^X] = \int_{-\infty}^{\infty} e^x \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) dx.$$

We use the same method as in Example 5. Gathering the exponential terms together, we get

$$\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} \exp\left(x - \frac{(x-\mu)^2}{2\sigma^2}\right) dx.$$

Now, focusing on

$$\left(x - \frac{(x-\mu)^2}{2\sigma^2}\right)$$

and completing the square gives

$$\begin{aligned}
-\frac{x^2 + x(-2\mu - 2\sigma^2) + \mu^2}{2\sigma^2} &= -\frac{x^2 - 2x(\mu+\sigma^2) + (\mu+\sigma^2)^2 - (\mu+\sigma^2)^2 + \mu^2}{2\sigma^2} \\
&= -\frac{(x-(\mu+\sigma^2))^2 - (\mu+\sigma^2)^2 + \mu^2}{2\sigma^2}.
\end{aligned}$$

Let $\hat{\mu} = \mu + \sigma^2$; then the integral becomes

$$\frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{\infty} \exp\left(-\frac{(x-\hat{\mu})^2}{2\sigma^2}\right)\exp\left(\frac{(\mu+\sigma^2)^2-\mu^2}{2\sigma^2}\right)\,dx.$$

The second exponential term doesn't depend on x, so we have

$$\exp\left(\frac{(\mu+\sigma^2)^2-\mu^2}{2\sigma^2}\right)\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}\sigma}\exp\left(-\frac{(x-\hat{\mu})^2}{2\sigma^2}\right)\,dx.$$

The integral is now the integral of the p.d.f. of a $N(\hat{\mu}, \sigma^2)$ random variable, so is 1. Hence the value of our integral is $\exp\left(\frac{(\mu+\sigma^2)^2-\mu^2}{2\sigma^2}\right)$ which rearranges to $\exp(\mu + \sigma^2/2)$ as required.

Challenge Questions

2.10 Let X be a random variable with a continuous distribution, and a strictly increasing distribution function F. Show that $F(X)$ has a uniform distribution on $(0, 1)$.

Suggest how we might use this result to simulate samples from standard distributions.

Solution. Since F is strictly increasing, it has an inverse function F^{-1}. For $x \in (0,1)$, we have

$$\mathbb{P}[F(X) \le x] = \mathbb{P}[X \le F^{-1}(x)] = F(F^{-1}(x)) = x.$$

Hence, $F(X)$ has the uniform distribution on $(0, 1)$.

We write this as $U = F(X)$, where U is uniform on $(0, 1)$. Therefore, $F^{-1}(U) = X$. Consequently, if we can simulate uniform random variables, and calculate $F^{-1}(x)$ for given x, we can simulate X as $F^{-1}(U)$.

In fact, this is a very common way of simulating random variables. Recall that a distribution function F is not necessarily strictly increasing, but it is necessarily *non-strictly* increasing. With some care, it is possible to extend this result to cover the general case. For many standard distributions, F^{-1} can be computed explicitly.

2.11 Prove that $\Gamma(\frac{1}{2}) = \sqrt{\pi}$.

Hint. Thanks to the normal distribution, you know that $\int_{-\infty}^{\infty} e^{-x^2/2}\,dx = \sqrt{2\pi}$.